Journal of Consulting and Clinical Psychology
2013, Vol. 81, No. 4, 722–736

0022-006X/13/$12.00 DOI: 10.1037/a0033004

Are the Parts as Good as the Whole? A Meta-Analysis of Component Treatment Studies

Erin C. Bell
University of Southern Mississippi

David K. Marcus
Washington State University

James K. Goodlad
University of Southern Mississippi

Objective: Component studies compare standard treatments to treatments with added components or dismantled components. A previous meta-analysis (Ahn & Wampold, 2001) failed to find any differences in outcome between treatments with more components and those with fewer components, leading the authors to conclude that common factors and not specific ingredients account for therapeutic change. ***Method:*** The current random effects model meta-analysis of psychotherapy component studies conducted between 1980 and 2010 included more than 3 times as many studies as Ahn and Wampold's (2001) meta-analysis ($k = 66$). Unlike the previous meta-analysis, this study conducted separate meta-analyses for additive and dismantling studies and also examined treatment outcomes at follow-up. ***Results:*** For the dismantling studies, there were no significant differences between the full treatments and the dismantled treatments. For the additive studies, the treatment with the added component yielded a small, but significant, effect at completion ($d = 0.14$) and a slightly larger effect at follow-up ($d = 0.28$), but only for the specific problems that were targeted for treatment. Despite the diversity of populations studied, problems treated, and treatments examined, there was little heterogeneity among the results of these studies. ***Conclusion:*** These findings suggest that added specific ingredients may contribute modestly to treatment outcomes.

Keywords: component studies, dismantling studies, additive studies, psychotherapy research, meta-analysis

In the years since the American Psychological Association's Division 12 (Clinical Psychology) Task Force on Promotion and Dissemination of Psychological Procedures (1995) published criteria for identifying empirically validated treatments, there has been considerable controversy over the contribution of specific treatment factors to psychotherapy outcome. Much of this controversy has focused on the findings from the so-called "Dodo Bird" meta-analyses (e.g., Luborsky, Singer, & Luborsky, 1975; Wampold et al., 1997), which failed to find evidence of differential outcomes when bona fide treatments are compared to one another. Although the results from these meta-analyses have been relatively consistent and a "meta-meta-analysis" of this literature (Luborsky et al., 2002) found support for the Dodo Bird hypothesis of treatment equivalence, numerous psychotherapy researchers have questioned this methodology and the conclusions drawn from these studies (e.g., Chambless, 2002; Crits-Christoph, 1997; Hunsley & Di Giulio, 2002).

Component studies (dismantling or additive) may provide a more direct method for identifying whether specific active ingredients in psychotherapy contribute to differential outcomes. In a dismantling design, at least one element of the treatment is removed and the full treatment is compared to this dismantled treatment. In additive designs, an additional component is added to an existing treatment to examine whether the addition improves the outcome. If the dismantled or added component is an active ingredient, then the condition with fewer components should yield less successful outcomes (e.g., Borkovec & Castonguay, 1998; Lang, 1969). This research design was first used to identify the active ingredients in systematic desensitization (e.g., Lang, 1969). More recently, Jacobson et al.'s (1996) influential dismantling study of cognitive-behavioral therapy (CBT) for depression compared the full package of CBT, to behavioral activation (BA) plus modification of automatic thoughts, to BA alone. Because this study failed to find differences among the three treatment groups, its findings were interpreted as indicating that BA was as effective as CBT, and the study contributed to a resurgence of interest in behavioral treatments for depression (Dimidjian, Martell, Addis, & Herman-Dunn, 2008).

Although component designs have a number of potential advantages over treatment comparison studies or treatment versus con-

This article was published Online First May 20, 2013.

Erin C. Bell, Department of Psychology, University of Southern Mississippi; David K. Marcus, Department of Psychology, Washington State University; James K. Goodlad, Department of Psychology, University of Southern Mississippi.

This work was based on Erin C. Bell's master's thesis at the University of Southern Mississippi.

Correspondence concerning this article should be addressed to David K. Marcus, Department of Psychology, Washington State University, P.O. Box 644820, Johnson Tower 233, Pullman, WA 99164-4820. E-mail: david.marcus@wsu.edu

trol studies, these designs have some limitations. Unlike in studies that include a common factors only control, null results in component studies are ambiguous with respect to the question of specific versus common factors, because both treatments include specific treatment components. Whereas significant effects in component studies directly support the efficacy of a specific factor, null results do not directly address the issue of specific versus common factors because there is no group that received only common treatment components. A second limitation is that it is possible that the dismantled component could interact synergistically with the other components of the treatment yielding results that overestimate the main effect of the dismantled component. In other words, removing a component might not only remove the main effect for that component, but also reduce the efficacy of the treatment by removing its interaction with other components. Component designs may also underestimate the contributions of the component. Rehm (2009) suggested that because much improvement typically occurs in the early stages of therapy, whichever component is presented first will appear to be the most effective. Thus, the dismantled component (which is never introduced) is likely to appear unnecessary.

Component studies are also likely to be statistically underpowered (Kazdin & Whitley, 2003) to detect the relatively small effect sizes that are likely to occur with these types of designs. Component studies compare a potentially effective treatment (the partial treatment) to the full treatment, which may be more effective. The ingredient that is removed or added is not expected to be the only active ingredient, or else the study would not be a component study, but instead would be a treatment versus placebo control study. The average effect size when an active treatment is compared to a placebo is $d = 0.48$ (Lambert & Bergin, 1994), so component studies should yield more modest effect sizes, even if an important treatment component is omitted from the partial condition. Therefore, a two-group component study with a presumed effect size of .24 (half the treatment vs. placebo effect size) would require over 250 patients in each condition to have a power of .80. Even Kazdin and Whitley's (2003) higher estimate of an effect size of .45 for additive design studies would require 78 patients in each condition. In contrast, the average sample size for the studies included in the present meta-analysis was 23 participants in each condition, which would require a large effect size of .84 to have a power of .80. Therefore, it is likely that many component studies will fail to yield significant group differences even when the omitted or added component is therapeutically efficacious.

Even if most individual component studies are underpowered, meta-analysis may provide a means for addressing the general question of whether these studies provide evidence supporting the role of specific factors in psychotherapy. Ahn and Wampold (2001) conducted such a meta-analysis and concluded that component studies provide no evidence in support of the efficacy of specific ingredients. Ahn and Wampold collected 20 articles reporting 21 component studies published in four journals (*Behaviour Research and Therapy*, *Behavior Therapy*, *Journal of Consulting and Clinical Psychology*, and *Journal of Counseling Psychology*) between 1990 and 1999. These studies yielded 27 treatment comparisons with an average effect size of $d = -0.20$, suggesting that the partial treatments may have outperformed the full packages, but this average effect size was not statistically significant. They did not examine potential moderators because the effect sizes in this meta-analysis were homogeneous.

There is, however, reason to believe that Ahn and Wampold's (2001) meta-analysis may not be the last word on this issue. Additional component studies have been published in the 12 years since their meta-analysis was accepted for publication. Also, although Ahn and Wampold justified their decision to limit their search to four journals based on the difficulty of conducting a keyword search using methodological terms, this approach omitted studies published in other journals. Additionally, given the limited number of studies in their sample, Ahn and Wampold combined the results from additive and dismantling studies. These two methodologies may involve different assumptions and may yield different results. Dismantling studies may be conducted to test whether certain components of the treatment are necessary, potentially resulting in more streamlined therapies. For example, 10 of the studies in our meta-analysis dismantled eye movement desensitization and reprocessing therapy (EMDR), and many of these studies appeared to have been designed to debunk the efficacy of lateral eye movements. In contrast, additive studies are often conducted by researchers who believe that the additional component will improve outcomes (i.e., why add a component if you believe it will be inert?). With a larger set of studies, separate meta-analyses for additive and dismantling component studies can be conducted. Furthermore, Ahn and Wampold compared treatment outcomes at termination, but did not examine follow-up outcomes. Finally, Ahn and Wampold combined both outcome measures that assessed the problems targeted for treatment and the nonspecific outcomes measures. This analytic strategy may attenuate treatment differences (Crits-Christoph, 1997). For example, when comparing CBT to other psychotherapies, Tolin (2010) found apparently larger effect sizes favoring CBT for measures of primary symptoms, global symptoms, and general functioning than for measures of self-concept and social adjustment.

The current meta-analysis used a larger set of component studies than those included in Ahn and Wampold's (2001) meta-analysis to examine whether the full treatment was superior to the partial treatment in dismantling studies, and whether adding a component improved outcomes in additive studies. For both sets of studies, we examined outcomes at both termination and follow-up. Additionally, for both sets of studies at both time points we examined both targeted treatment outcomes and outcomes assessed using other non-specific measures. Thus, the present study yielded eight effect sizes. Finally, even if adding components modestly improves the outcomes, there is the potential limitation that additional treatment components could increase attrition by requiring additional effort on the part of the clients (Kazdin & Whitley, 2003). Conversely, a dismantled treatment may reduce attrition because it makes fewer demands on clients than the full treatment. To address this issue, we also examined attrition rates across the various treatment groups at both completion and follow-up.

Method

Identification of Studies

The current meta-analysis generally adopted the same inclusion criteria used by Ahn and Wampold (2001). Specifically, the study had to (a) involve a psychological treatment for a particular dis-

order or condition, (b) include the removal or addition of at least one component of the treatment that was presumed to be efficacious, and (c) include information to calculate an effect size. Additionally, the dismantled condition had to be a viable treatment (i.e., not simply a placebo control). Studies also had to have been written in English and published in peer-reviewed journals. Unlike Ahn and Wampold, we included studies that did not specify that the therapists had at least a master's degree; instead, studies were included as long as the therapy was provided by a graduate level student or a therapist supervised by a professional.

Because relying on keyword searches to identify component studies is likely to overlook many relevant studies (Ahn & Wampold, 2001), we used a variety of strategies to locate studies for the current meta-analysis. We identified records by including the articles from Ahn and Wampold's (2001) meta-analysis, reviewing the tables of contents from 2000 to 2010 for the four journals identified in Ahn and Wampold's meta-analysis, and reviewing the reference sections from various articles and chapters likely to discuss component studies. We also searched PsycINFO and Medline using the combined search terms "dismantling" and "treatment" for all studies published between 1980 and 2010, which yielded 148 possible references (see Figure 1). Ahn and Wampold's meta-analysis contributed 21 studies to the meta-analysis; the table of contents review yielded an additional 24 studies; Longmore and Worrell's (2007) narrative review of whether cognitive interventions are necessary cited three studies not previously identified; Herbert et al.'s (2000) critique of EMDR included seven additional dismantling studies; the literature reviews in Barlow (2008) added six more studies; Rehm's (2009) commentary on component studies cited a component study he had conducted; and the database search yielded five studies that had not been identified from the previous sources. Thus, the meta-analysis included 66 studies that included 3,244 participants. The studies in the meta-analysis are summarized in Table 1.

Coding and Analyses

The effect sizes were computed as Cohen's *d* (the mean of the fewer-component-group minus the mean of the more-component-group divided by the pooled standard deviation). To be consistent with Ahn and Wampold's (2001) meta-analysis, a positive *d* indicates that the larger treatment package was superior to the smaller condition (measures indicating better functioning were reverse scored). When a study provided multiple targeted outcome measures of the treatment problem, an effect size was calculated for each measure and then they were averaged. This procedure was also used when studies reported multiple non-specific outcome measures. We classified those measures of the disorder being treated as the specific measures (e.g., depression measures were specific for depression treatment studies, but not for panic disorder studies). Usually, this classification followed the original study's identification of measures as primary or secondary, but if a study classified a targeted measure as secondary because another measure was considered the "gold standard," we still classified that measure as a targeted measure. For example, Bryant et al.'s (2008) post-traumatic stress disorder (PTSD) study explicitly listed the Impact of Events Scale as a secondary measure, but because it is a specific measure of trauma impact, we included it with the

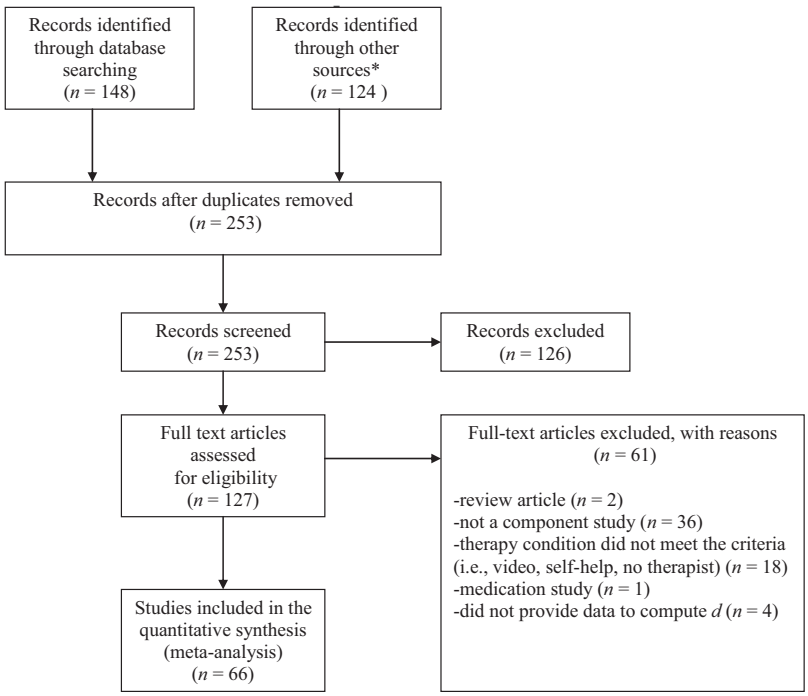

Figure 1. Flow chart of study selection. *Other sources included a hand review of the table of contents of *Behaviour Research and Therapy*, *Behavior Therapy*, the *Journal of Consulting and Clinical Psychology*, and the *Journal of Counseling Psychology* from 2000 to 2010; the studies used in Ahn and Wampold (2001); and the reference sections of various reviews.

Table 1
Studies in the Meta-Analysis

Additive studies				Targeted outcomes				Non-targeted outcomes			
Study	Disorder	Standard treatment (n^a)	Added component (n)	*d* at PT	*SE* at PT	*d* at FU	*SE* at FU	*d* at PT	*SE* at PT	*d* at FU	*SE* at FU
Appelbaum et al. (1990)	Tension headache	PMR (17)	CT (16)	−.45	.35						
Baucom et al. (1990)	Marital discord	BMT (12)	CR (12) EET (12) EET + CR (12)	−.24[b]	.41	−.01[b]	.41	−.29[b]	.41	.03[b]	.41
Blanchard et al. (1990)	Tension headache	PMR (19)	CT (16)	−.12	.34						
Bryant et al. (2003)	PTSD	IE (20)	CR (20)	.25	.32	.36	.32	.29	.32	.21	.32
Bryant et al. (2005)	Acute stress	CBT (33)	Hypnosis (30)	.04	.25	.02	.25	.19	.25	.10	.25
Bryant et al. (2008)	PTSD	IE + In vivo EX (31)	CT (28)	.34	.26	.81	.32	.33	.26	.62	.32
Butler et al. (1984)	Social phobia	EX (15)	CT (15)			1.01	.40				
Dadds & McHugh (1992)	Child conduct problems	CMT (11)	Ally (11)	.19	.43	.17	.43	.32	.43	.27	.43
de Jong et al. (2000)	Spider phobia	EX (16)	Counter-conditioning (18)	−.25	.34	−.24	.41	−.27	.35	−.23	.41
de Zwaan et al. (2005)	Binge eating	BT (VLCD) (35)	CBT (36)	−.17	.24						
Deffenbacher et al. (2002)	Angry drivers	Relaxation coping skills (16)	Cognitive skills (17)	.35	.35	.02	.35				
Deffenbacher & Stark (1992)	Anger	Relaxation coping skills (19)	Cognitive skills (16)	−.23	.34	−.19	.39	.01	.34	.02	.39
Emmelkamp & Beens (1991)	OCD	EX (11)	CT (10)	.32	.44	1.15	.47	.26	.44	.20	.44
Foa et al. (2005)	PTSD	PE (79)	CR (74)	.08	.16	.06	.21	.10	.16	.07	.21
Foa & Rauch (2004)	PTSD	PE (27)	CR (27)	.22	.27	.10	.33				
Grunes et al. (2001)	OCD	ERP (14)	FI (14)	.42	.38	.46	.38	−.09	.38	.04	.38
Halford et al. (1993)	Marital discord	BMT (13)	CR, affect exploration, and generalization training (13)	−.31	.39	.06	.39				
Hiss et al. (1994)	OCD	EX (10)	Relapse prevention (8)			.89	.50			.41	.48
Kazdin & Whitley (2003)	Child conduct	PSST + PMT (70)	PPS (57)	.32	.18			.43	.18		
Koch et al. (2004)	Small animal phobia	EX (20)	CBT (20)	.48	.32	.74	.33	.21	.32	.51	.32
Mattick & Peters (1988)	Social phobia	EX (23)	CR (21)	.02	.30	.29	.30	−.20	.30	−.15	.30
McKay et al. (2010)	Cocaine dependence	CM (24) RP (20)	RP (24) CM (24)	.24	.30	.74	.31				
Mehta (1990)	OCD	BT alone (15)	FI (15)	.92	.38	1.48	.41	.83	.38	1.23	.40
Nicholas et al. (1991)	Chronic back pain	BT (9)	PMR (8)	−.79	.50	.47	.65	.54	.49	1.41	.72
Nicholas et al. (1991)	Chronic back pain	CT (7)	PMR (8)	1.57	.59	1.32	.70	.65	.53	.88	.66
Paunovic & Öst (2001)	PTSD	EX (9)	CBT (7)	.30	.51	.15	.50	−.22	.51	−.19	.51
Petry et al. (2008)	Gambling	Motivational enhancement (55)	CBT (40)	.31	.21	.28	.23				
Porzelius et al. (1995)	Binge eating	OBET (21)	Advanced CBT and coping skills (25)	−.07	.30	.07	.32	−.32	.30	−.17	.32
Propst et al. (1992)	Depression	CBT (10)	Religious component (10)	.36	.45	.19	.45	1.07	.48	−.29	.45
Radojevic et al. (1992)	Arthritis	BT (14)	Social support (15)	.22	.37	.07	.37	.23	.37	.24	.37
Rohan et al. (2007)	SAD	CBT (15)	Light therapy (16)	.47[b]	.37	.17[b]	.40				
Rosen et al. (1990)	Body image	CBT (11)	Perception training (13)	−.18	.41	−.18	.41	−.12	.41	−.09	.41
Sanders & McFarland (2000)	Family behavior	Behavioral family intervention (24)	Cognitive component (23)	.06	.29	.17	.32	−.35	.29	−.15	.32
Schmiege et al. (2009)	HIV/STD risk	GPI (154)	GMET (163)	.24	.11	.10	.14				
Thackwray et al. (1993)	Bulimia	BT (13)	CT (13)	−.76	.41	.26	.39	−.62	.40	−.18	.39
Webster-Stratton (1994)	Parenting	GDVM (39)	ADV (38)	.12	.23	.14	.30	.08	.23	.04	.30

(*table continues*)

Table 1 (*continued*)

Dismantling studies				Targeted outcomes				Non-targeted outcomes			
Studies	Disorder	Standard treatment	Component removed	*d* at PT	*SE* at PT	*d* at FU	*SE* at FU	*d* at PT	*SE* at PT	*d* at FU	*SE* at FU
Barlow et al. (1989)	Panic disorder	PMR + EX + CR (16)	PMR (15) EX + CR (10)	−.18[b]	.38	−.38[b]	.54	.02	.38	.61	.55
Barlow et al. (1992); FU Craske et al. (1991)	GAD	CR + PMR (11)	PMR (13) CR (10)	−.26[b]	.43			.05	.42		
Bauman & Melnyk (1994)	Test anxiety	EMDR (15)	Eye-movement (15)	−.03	.37			−.50	.37		
Borkovec et al. (2002)	GAD	CT + SCD (23)	CT (23) SCD (23)	.16[b]	.30	.27[b]	.30	.20	.30	.20	.30
Borkovec & Costello (1993)	GAD	CBT (19)	AR (18)	.19	.33	−.04	.34	.13	.33	.24	.34
Cusack & Spates (1999)	PTSD	EMDR (11)	Eye-movement (11)	.00	.43			.66	.44		
Devilly et al. (1998)	PTSD	EMDR (13)	Eye-movement (12)	.33	.40	.10	.46	.39	.40		
Dunn et al. (1996)	Anxiety	EMDR (14)	Eye-movement (14)	.00	.38						
Feske & Goldstein (1997)	Panic disorder	EMDR (18)	Eye-movement (18)	.00	.33	.07	.38	.28	.33	−.35	.38
Flessner et al. (2005)	Nail biting	SHR (18)	Social support (20)	.14	.33						
Foa et al. (1984)	OCD	EX + RP (11)	EX (9) RP (12)	.75[b]	.45	.40[b]	.44	.59	.44	.15	.43
Foley & Spates (1995)	Public speaking phobia	EMDR (10)	Eye-movement (10) EX + MS (10)	−.04[b]	.45						
Gosselin & Matthews (1995)	Test anxiety	EMDR (21)	Eye-movement (20)	.12	.31						
Hope et al. (1995)	Social phobia	CBT (13)	EX (10)	−.25	.42	−.28	.49	.13	.42	−.24	.49
Jacobson et al. (1996)	Depression	CBT (49)	BA + AT (55) BA (42)	−.06[b]	.20	−.02[b]	.21				
Marks et al. (1998)	PTSD	EX + CR (20)	EX (19) CR (20)	−.36[b]	.32	.18[b]	.43	.10	.32	.69	.44
Mattick et al. (1989)	Social phobia	CBT (11)	CR (11) EX (11)	.32[b]	.43	−.30[b]	.46				
Nezu & Perri (1989)	Depression	PST (14)	Problem- orientation component (14)	1.29	.42	1.21	.41				
Öst et al. (2004)	Panic disorder	CBT (34)	EX (29)	−.05	.25	−.03	.25				
Öst et al. (1991)	Blood phobia	ATP (10)	Tension (10) EX (10)	.10[b]	.45	.24[b]	.45				
Pitman et al. (1996)	PTSD	EMDR (16)	Eye-movement (14)	−.23	.49						
Rehm (1981)	Depression	SM + SE + SR (12)	SM + SR (11) SM + SE (12)	.06[b]	.41			−.22	.41		
Resick et al. (2008)	PTSD	CPT package (42)	CPT-writing (38) Written accounts (38)	.12[b]	.22	.16[b]	.22	.01	.22	−.05	.22

(*table continues*)

Table 1 (*continued*)

Dismantling studies				Targeted outcomes				Non-targeted outcomes			
Studies	Disorder	Standard treatment	Component removed	*d* at PT	*SE* at PT	*d* at FU	*SE* at FU	*d* at PT	*SE* at PT	*d* at FU	*SE* at FU
Roehrig et al. (2006)	Body image	Counter attitudinal therapy + EX (30)	Counter attitudinal component (36)	−.47	.25	−.22	.29	−.08	.25	.42	.29
Sanderson & Carpenter (1992)	Specific phobia	EMDR (32)	Eye-movements (30)	.09	.25						
Schmidt et al. (2000)	Panic disorder	CBT + Breathing retraining (32)	CBT (21)	−.37	.28	−.40	.32				
Taylor et al. (1997)	Social phobia	CT + EX (22)	EX (22)	.34	.30	.31	.33	.62	.31	.18	.33
Taylor et al. (2003)	PTSD	EMDR (15)	Eye-movements (15)	−.63	.37			−.17	.37		
Walters et al. (2009)	Alcohol problems	MI + Feedback (70)	Feedback (59) MI (57)	.02[b]	.18	.18[b]	.18				
Williams & Falbo (1996)	Panic disorder	CBT (11)	BT (13) CT (10)	−.07[b]	.42	−.15[b]	.42	.21	.42	.14	.42

Note. PT = post-treatment; FU = follow-up. Disorder: GAD = generalized anxiety disorder; PTSD = post-traumatic stress disorder; OCD = obsessive-compulsive disorder; SAD = seasonal affective disorder; HIV = human immunodeficiency virus; STD = sexually transmitted disease. Component group: CT = cognitive therapy; PMR = progressive muscle relaxation; EX = exposure; CR = cognitive restructuring; BMT = behavioral marital therapy; EET = emotional expressiveness training; EMDR = eye movement desensitization reprocessing; SCD = self-control desensitization; AR = applied relaxation; CBT = cognitive-behavioral therapy; IE = imaginal exposure; CMT = child management training; BT = behavior therapy; VLCD = very low calorie diet; SHR = simplified habit reversal; RP = response prevention; PE = prolonged exposure; EX + MS = exposure and movement to sound; ERP = exposure and response prevention; FI = family involvement; BA = behavioral activation; AT = automatic thoughts; PSST = problem solving skills training; PMT = parent management training; PPS = parent problem solving; CM = contingency management; PST = problem solving therapy; ATP = applied tension package; CPT = cognitive processing therapy; OBET = obese binge Eating treatment; SM = self-monitoring; SE = self-evaluation; SR = self-reinforcement; GPI = theory-based sexual risk reduction intervention; GMET = group-based motivational enhancement therapy; MI = motivational interviewing; GDVM = videotaped parent skills training program; ADV = cognitive training social learning program.
[a] These values are the sample sizes at treatment completion. [b] The average *d* for the study.

targeted measures. To assess the reliability of the decisions about targeted versus non-targeted measures, a second coder judged these measures for a random sample of 12 studies involving 88 outcomes measures. The two raters agreed about the classification of 87 of these 88 measures. For the disagreement, the decision made by the first author was used.

To combine the results across studies, the effect sizes were weighted by the inverse variance so that studies with larger samples had greater weight when computing the average *d*. We used a random effects model to combine the effect sizes across studies. A random effects model is the appropriate model for this meta-analysis, because unlike a fixed effect model, it does not assume that variation across studies is solely due to chance. Thus, there is not one overall effect but that the true treatment effect varies across studies because of actual differences in the study characteristics, such as the treatment provided and the population being treated (Riley, Higgins, & Deeks, 2011). Consequently, the average *d* value in a random effects model is "the average effect rather than the common effect" (Riley et al., 2011, p. 965).

Fifteen of the studies included in this meta-analysis involved two partial conditions (e.g., a full CBT package was compared to a treatment in which schema therapy was removed and to another condition in which both schema therapy and cognitive restructuring were removed), and one study had three partial conditions. In these instances, Ahn and Wampold (2001) treated each comparison as an independent effect size, which violated the assumption of independence (i.e., the same treatment comparison group contributed to multiple effect sizes). Instead, we averaged each of these within-study effect sizes so that the independence assumption was not violated. With respect to follow-up data, some studies reported outcomes at multiple follow-up periods. Because 6 months was the modal follow-up time (28 of the 51 studies that reported follow-up data reported a 6-month follow-up), when studies reported several follow-up data points, the point closest to 6 months was used to compute the effect size.

To determine whether effect sizes were from a single population, a Q test was calculated. A statistically significant Q value indicates that the effect sizes were heterogeneous. I^2 "describes the percentage of total variation across studies that is due to heterogeneity rather than chance" (Higgins, Thompson, Deeks, & Altman, 2003, p. 558). An I^2 value of 0 indicates no heterogeneity. I^2 values of 25% or less are considered low levels of heterogeneity (Higgins et al., 2003). To investigate potential publication bias, we generated funnel plots with the effect sizes plotted along the *x*-axis and the corresponding standard errors on the *y*-axis (largest values at the bottom), examined the funnel plot asymmetry using a regression test (Egger, Smith, Schneider, & Minder, 1997), and conducted trim and fill analyses (Duval & Tweedie, 2000). Funnel plots that are symmetrical with the studies with effect sizes closest to the average effect size at the top (i.e., smallest standard errors) suggest an absence of publication bias. A significant regression test indicates that the funnel plot is asymmetrical and suggests publication bias. The trim and fill procedure examines the funnel plot for asymmetry and adds presumed missing values to make the plot symmetrical. The average effect size is then recalculated with these missing values. A substantial change in the average effect size suggests publication bias.

In addition to computing effect sizes, each study was coded for sample size, whether the study used an additive or dismantling

design, intervention type (e.g., CBT), client age, client gender (percent male), whether the conditions had an equal number of sessions, and year of publication. All of the analyses were conducted using Lipsey and Wilson's (2001) SPSS statistical programs and Viechtbauer's (2010) "metafor" package for R.

Results

Description of the Included Studies

All but one of the studies reported that participants were randomly assigned to the treatment conditions (Sanderson & Carpenter, 1992, failed to indicate whether they used random assignment). Thirty-six of the studies used an additive design, and 30 used a dismantling design. The average number of sessions provided was 9.6 (SD = 6.5), ranging from eight studies that examined a single-session treatment to five studies that examined treatments ranging from 20 to 29 sessions. The treatment sessions ranged in length from 45 to 180 min (M = 83.7, SD = 28.8). Most (k = 62) of the studies treated adult samples, with one study of adolescents, two studies of primary school children, and one family study that included both mothers and children. Ten of the studies included exclusively female participants, one study was exclusively male, and the remainder (k = 55) were mixed. Across studies, the samples were 64.2% female.

Additive Design Studies

Table 2 summarizes the findings from meta-analyses. Thirty-four of the studies that used an additive design provided data for computing effect sizes at completion (data from two studies could only be used to compute effect sizes at follow-up). There was some evidence that the treatment packages with the added components were slightly superior to the standard treatments at the completion of therapy for addressing the targeted problems. The average d across these 34 studies was 0.14, which was small, but statistically significant, 95% CI [.03, .24], Z = 2.61, p = .009. There was little heterogeneity among these studies, $Q(33)$ = 36.41, p = .31, I^2 = 9.37%. There was also little evidence suggesting publication bias.

Table 2
Result of the Meta-Analyses

Design	k	d	95% CI	SE	Q	I^2 (%)
Additive						
Termination						
Targeted	34	.14**	[.03, .24]	.052	36.41	9.37
Non-targeted	24	.12	[−.02, .25]	.069	26.17	12.10
Follow-up						
Targeted	32	.28***	[.13, .38]	.065	37.15	16.54
Non-targeted	24	.14	[−.00, .28]	.073	22.28	0.00
Dismantling						
Termination						
Targeted	30	.01	[−.11, .12]	.058	27.11	0.00
Non-targeted	17	.12	[−.04, .28]	.083	11.22	0.00
Follow-up						
Targeted	19	.08	[−.07, .22]	.072	15.65	0.00
Non-targeted	11	.15	[−.05, .36]	.104	6.37	0.00

Note. CI = confidence interval.
** $p < .01$. *** $p < .0001$.

The funnel plot appeared symmetrical (see Figure 2a), and the regression test for funnel plot asymmetry was not significant (z = −0.61, p = .54). The trim and fill procedure only imputed two missing studies on the right side of the graph, and adding these two data points had little effect on the effect size estimate (d = 0.16).

Only 24 of the studies that used an additive design reported results involving non-targeted outcome measures. Although similar in magnitude to the outcomes assessed with targeted measures, the average effect size for these non-targeted outcomes was not significant, d = 0.12 (95% CI [−.02, .25], Z = 1.67, p = .09). There was little heterogeneity among these studies, $Q(23)$ = 26.17, p = .29, I^2 = 12.10%. The funnel plot appeared symmetrical,[1] the regression test indicated that there was no evidence of asymmetry (z = 0.04, p = .97), and the trim and fill procedure indicated that there were no missing studies.

Thirty-two additive studies reported targeted outcomes at follow-up. The average d was 0.28 (95% CI [.15, .41], Z = 4.29, $p < .0001$). There was little heterogeneity among these studies, $Q(31)$ = 37.15, p = .21, I^2 = 16.54%. The funnel plot was generally symmetrical, although there was one outlier (see Figure 2b), and the regression test was significant (z = 1.99, p = .047). The trim and fill test indicated that there were no missing studies. Rerunning the analysis omitting the outlier yielded a slightly smaller estimate, d = 0.24 (95% CI [.12, .35], Z = 4.10, $p < .001$), and the regression test for asymmetry was no longer significant (z = 1.76, p = .08).

The 24 additive studies that measured non-targeted outcomes at follow-up yielded an average effect size of d = 0.14 (95% CI [−.00, .28], Z = 1.90, p = .058). These studies were homogeneous, $Q(23)$ = 22.28, p = .50, I^2 = .00%. There was no evidence of publication bias: The funnel plot was symmetrical, the trim and fill test indicated that there were no missing studies, and the regression test was not significant (z = 1.26, p = .21).

The apparently larger improvements at follow-up for the targeted outcomes (.28 or .24 omitting the outlier) compared to at termination (.14) raises the possibility that those studies that reported both termination and follow-up targeted outcomes may have differed from those that did not. To examine this possibility, we reran the analysis for the targeted outcomes for the 30 additive studies that included both termination and follow-up data. Although the average effect size increased slightly at termination (d = 0.16, p = .003) and decreased slightly at follow-up (d = 0.25, $p < .001$), it seems unlikely that this methodological issue fully accounts for the apparently greater differences found at follow-up compared to termination.

These results raise the question of whether the greater difference between the treatment conditions at follow-up for targeted outcomes was because of increased improvement following treatment in the added treatment condition, or because the treatment gains for those in the standard treatment condition dissipate with time. To address this question, we computed standardized mean gain scores (the mean at follow-up minus the mean at termination divided by the pooled standard deviation) examining the change from termination to follow-up for both types of treatment. Two studies that reported both termination and follow-up data did not provide

[1] Funnel plots for the analyses of the non-targeted outcomes and all of the dismantling studies are available on request.

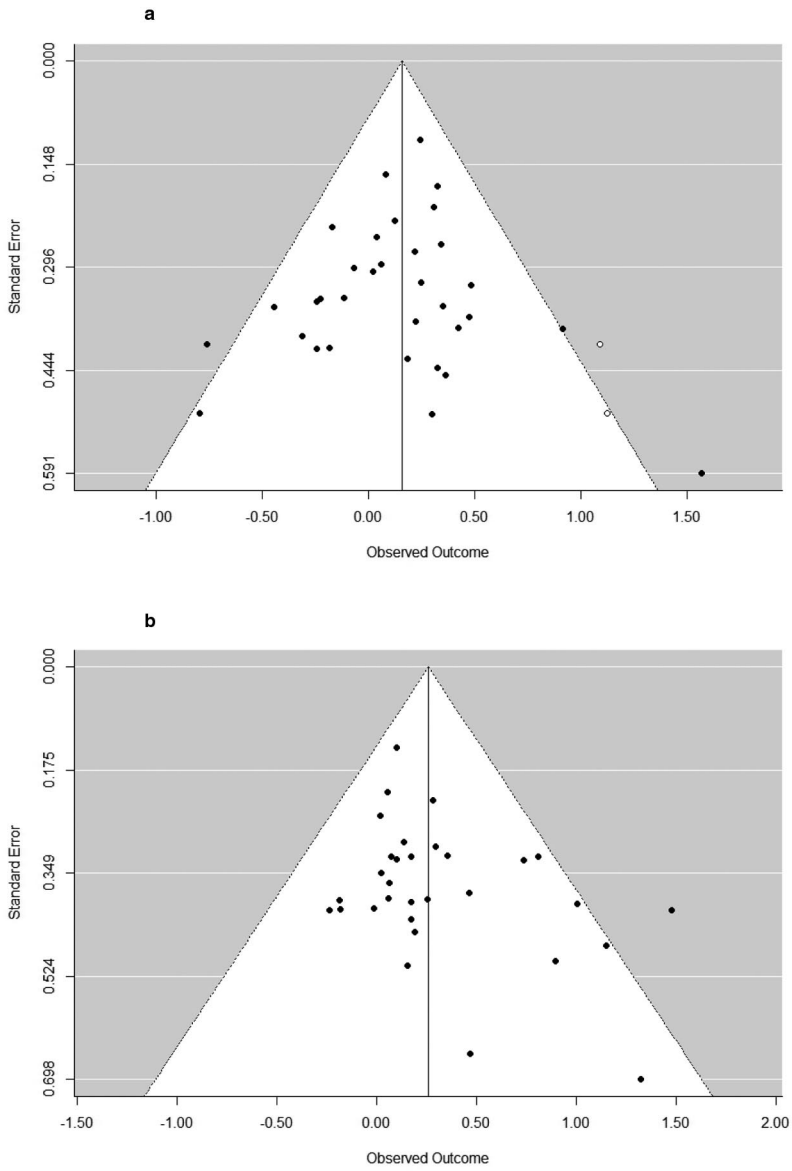

Figure 2. Funnel plot including the trim and fill procedure. Black circles represent the actual data, and the white circles represent imputed missing values. Figure 2a is the targeted outcomes for additive studies at termination, and Figure 2b is the targeted outcomes for additive studies at follow-up.

means and standard deviations, resulting in 28 studies for these analyses. On average, the participants in the added treatments continued to improve from termination to follow-up, $d = 0.20$ (95% CI [.08, .31], $Z = 3.42$, $p = .0006$). There was little evidence of continued improvement in the standard treatments, $d = 0.06$ (95% CI [−.07, .18], $Z = 0.91$, $p = .36$). It appears that the increased difference between those who received an additional component and those who received the standard treatments at follow-up is probably due to continued improvement for those clients who received the additional component and not because those in the standard treatment got worse.

Consistent with the dose–effect relationship (Howard, Kopta, Krause, & Orlinsky, 1986), it is possible that the clients in the added component condition improved more simply because they received a larger dose of therapy. Eight of the studies administered the added component using extra sessions, and 26 of the additive studies kept the amount of therapy contact consistent across both the standard treatment and the added component condition. At termination, these two designs yielded almost identical effect sizes, with the added sessions design yielding an effect of .143, and the equal sessions design yielding an effect of .145. Only four studies that added sessions reported outcomes at follow-up and these studies yielded an average effect of .33 compared to an average effect of .25 in the 28 additive studies with an equal number of sessions across the conditions. The difference between these two effect sizes was not significant, $Q_B(1) = 0.28$, $p = .60$.

Dismantling Studies

Across the 30 dismantling studies, there was no evidence that full treatments yielded better targeted outcomes at termination, $d = 0.01$ (95% CI [−.11, .12], $Z = 0.11$, $p = .91$).[2] There were also no significant differences between the full treatment and the dismantled treatment when non-target outcomes were assessed, $d = 0.12$ (95% CI [−.04, .28], $k = 16$, $Z = 1.43$, $p = .15$). Both sets of data were homogeneous (see Table 2). For both analyses, the funnel plots were symmetrical, the trim and fill tests indicated that there were no missing studies, and the regression tests were not significant.

There were no significant differences between the full treatment and the dismantled treatment at follow-up for either the targeted ($d = 0.08$, 95% CI [−.07, .22], $k = 19$, $Z = 1.05$, $p = .29$) or the non-targeted ($d = 0.15$, 95% CI [−.05, .36], $k = 11$, $Z = 1.46$, $p = .14$) outcomes. There was little-to-no heterogeneity among either set of studies (see Table 2). For the targeted outcomes, the trim and fill procedure indicated that there were two missing studies on the right side of the funnel graph, and when these studies were imputed, the average d increased to 0.11 ($p = .11$). For the non-targeted outcomes, the trim and fill procedure imputed two missing studies on the left side of the plot, reducing d to 0.10 ($p = .32$). Neither regression test was significant.

Attrition

For both the additive and dismantling studies, we calculated the percentage of clients in each condition who (a) completed treatment, (b) were retained from treatment termination to follow-up, and (c) were retained from enrollment to follow-up. As can be seen in Table 3, there were only minor differences in attrition rates across the four groups, with studies generally reporting high rates of client retention.

Supplementary Analyses

To examine whether our search strategy could explain the difference between the current results and those reported by Ahn and Wampold (2001), we conducted a series of analyses using various subsets of the studies that were meta-analyzed (see Table 4). To be consistent with Ahn and Wampold, we combined the results from the additive studies with the results from the dismantling studies and also combined the targeted and non-targeted outcome measures. Like Ahn and Wampold, we did not find a significant difference between the full treatments and the partial treatments for the 21 studies they originally meta-analyzed, and there was no significant effect at follow-up. When we added the studies that were published between 2000 and 2010 in the four journals that were reviewed by Ahn and Wampold, there was a small but significant d (0.09) at termination and at follow-up ($d = 0.12$). Adding five more studies from these same journals that were published in the decade before the period reviewed by Ahn and Wampold yielded almost identical results, and adding all of the studies had little impact on the average effect sizes.

Table 3
Attrition Across the Study Designs and Treatments

Design	Treatment completion		Completion to follow-up		Enrollment to follow-up	
	k	%	*k*	%	*k*	%
Additive	33		29		28	
Standard treatment		87.3		92.1		79.5
With added component		88.0		89.6		77.8
Dismantling	28		20		20	
Standard treatment		89.0		87.5		76.9
Dismantled treatment		90.2		86.3		76.1

Discussion

The current meta-analysis partially replicated and extended the findings from Ahn and Wampold's (2001) pioneering meta-analysis of component studies. Ahn and Wampold failed to find evidence that full treatment packages outperformed partial treatment conditions at the completion of therapy. However, they combined additive and dismantling design studies and did not examine outcomes at follow-up. In the current meta-analysis, we separated the additive and dismantling studies. For the dismantling studies, there was little evidence that removing a therapy component had an adverse effect on the treatment outcomes, and this finding is consistent with Ahn and Wampold's conclusions. In contrast, among the additive studies, there was a small but significant average effect favoring treatments that included an added component at termination, and a somewhat larger effect at follow-up. These effects were only significant for targeted outcomes (or for the combination of targeted and non-targeted outcomes) and not when only non-targeted outcomes were assessed.

Although meta-analysis allows researchers to detect small effects, one challenge is how to interpret these modest effect sizes. Following Cohen's (1988) guidelines, the significant effect sizes of 0.14 and 0.28 (or 0.24) for additive studies at termination and follow-up would be considered small. Higgins and colleagues (Higgins, Thompson, & Spiegelhalter, 2009; Riley et al., 2011) have recently recommended that the 95% prediction interval, in this case the predicted range of a future additive study, can help with the interpretation of the results of a random effects meta-analysis. The 95% prediction interval for targeted outcomes at termination was −.06 to .34, and at follow-up it was −.05 to .61,[3] so that in most future additive studies patients who receive the added component are likely to improve more than those who receive the standard treatment. However, under some circumstances those who receive the added component may not show greater improvement and may even do slightly worse, even at follow-up. In contrast, the 95% prediction interval for the targeted outcomes for the dismantling studies was .00 to .01 at termination and .03 to .12 at follow-up, indicating that future dismantling studies are likely to find few to no differences between the full treatment and the dismantled condition.

Although an average effect size of 0.24 to 0.28 at follow-up is not large, it may not be trivial. Applying Rosenthal and Rubin's

[2] This small effect was not simply due to the inclusion of EMDR dismantling studies in the meta-analysis. The 10 EMDR dismantling studies had an average d of −0.02 ($p = .88$), and the other 20 dismantling studies had an average d of 0.01 ($p = .83$).

[3] However, because of the reduced heterogeneity, the 95% prediction interval at follow-up when the outlier was omitted was .12–.36.

Table 4
Results From Different Subsets of Studies

Source	Termination				Follow-up			
	k	*d*	*Q* (*df*)	I^2	*k*	*d*	*Q* (*df*)	I^2
1. Original Ahn and Wampold (2001) studies	21	−.08	14.21 (20)	0	18	.05	5.78 (17)	0
2. Studies from 1 plus the studies published in the journals Ahn and Wampold reviewed between 2000 and 2010	45	.09*	33.64 (44)	0	38	.12*	21.66 (37)	0
3. Studies from 1, 2, plus the studies published in the journals Ahn and Wampold reviewed between 1980 and 1989	50	.10*	44.49 (49)	0	44	.14**	34.31 (43)	0
4. All studies	64	.10**	55.56 (63)	0	51	.17***	45.87 (50)	0

* $p < .05$. ** $p < .01$. *** $p = .0001$.

(1982) binomial effect size display to an estimated effect of 0.24 ($r = .12$), indicates that there will be a 56% success rate at follow-up for those receiving the added component compared to a 44% success rate for those who receive the standard treatment (where success is defined as the median outcome score on the targeted measures at follow-up). Adding an additional component to an active treatment is likely to lead to a slightly improved outcome over time at least with respect to the targeted problem. If this component does not lead to increased attrition or significantly increased costs, it may be worth the effort. This finding does not call into question the importance of general factors in psychotherapy but does suggest that specific added components can make modest contributions to long-term outcomes.

In addition to separating the additive from the dismantling studies, we also ran separate analyses for the targeted and non-targeted outcomes. For the additive studies, the average effect sizes for the targeted and non-targeted outcomes were similar at termination, but at follow-up the average effect size for the targeted outcomes was twice as large as for the non-targeted outcomes (0.28 vs. 0.14). Advocates of specific factors in psychotherapy may interpret these results as supporting Crits-Christoph's (1997) point that including non-targeted outcomes in psychotherapy outcome meta-analyses is likely to dilute the apparent effectiveness of the specific treatment factor. Alternatively, those who are skeptical of the value of specific factors can argue that adding a component to therapy will at best only have a limited impact on the overall treatment outcome. How one interprets these results may be largely driven by one's assumptions about the aims and mechanisms of psychotherapy. If the aim of therapy is to reduce targeted symptoms, then adding these components can improve the treatment's long-term efficacy; however, if the aim is to improve the patient's quality of life, then these specific components may have more limited value.

Allegiance effects (e.g., Luborsky et al., 1999) may, at least partly, account for the differences between the additive and dismantling studies. Allegiance effects may not necessarily reflect researcher bias, but instead may reflect the researchers' empirically informed hypotheses (Leykin & DeRubeis, 2009; but see Munder, Gerger, Trelle, & Barth, 2011, for a meta-analysis suggesting that the allegiance effect is due to allegiance bias). Researchers rarely add additional components to established treatments unless they believe that there is a reasonable possibility that these new components will improve the efficacy of the treatment. Conversely, a frequent rationale for conducting dismantling studies is to identify treatment components that may be superfluous, whether they are eye movements in EMDR (e.g., Taylor, Thordarson, Fedoroff, Maxfield, & Lovell, 2003) or breathing retraining in CBT for panic disorder (Schmidt et al., 2000).

Instead of concluding, as Ahn and Wampold (2001) did, that component studies provide evidence that nonspecific factors are mainly responsible for therapeutic change, our findings suggest a more nuanced conclusion. First, even null results from component studies do not directly speak to the role of specific factors. To take an example from medicine, if a study found that chemotherapy plus radiation were no more effective than chemotherapy or radiation alone for treating certain cancers, these results would not indicate that non-specific factors were responsible for the effectiveness of cancer treatment. Second, despite some of the limitations inherent in drawing causal attributions from follow-up data (discussed below), it does appear that clients who received an added component in therapy continued to improve on the targeted outcomes somewhat more after therapy than those who received the standard treatment. Although, the data from the present meta-analysis cannot directly address the issue of mechanisms of change, our findings suggest various possibilities that may be examined in future research. It may be that treatment outcomes were relatively similar at the termination of therapy regardless of whether a component had been added, either because nonspecific factors were most influential at the time of termination or because whatever specific ingredients were provided in the standard condition were sufficient for initial change. Additional ingredients might have had a "sleeper effect" at follow-up because they may (a) provide clients with specific tools that increase their chances of improving once therapy is over (a possibility that is indirectly supported by the apparently larger effects for the targeted measures than for the more general measures), (b) give clients a broader range of options for dealing with their problems and thus increase their sense of agency (e.g., Bohart & Tallman, 1999), or (c) lead clients to report greater change to justify the greater effort that may have accompanied the treatment with the added component (i.e., through dissonance reduction). However, because most additive studies kept the amount of therapy constant across the conditions (i.e., clients who did not receive the added component still received the same number of sessions), and there were few differences between additive studies that kept the number of sessions constant and those that added sessions, these effects were

probably not simply due to clients receiving more treatment in the added component condition. Regardless of the mechanism, these findings underscore the importance of collecting and reporting follow-up outcomes when conducting psychotherapy component studies.

Even with a larger sample of studies drawn from a wider range of journals over a greater time span, the current study replicated Ahn and Wampold's (2001) finding that there was little heterogeneity across component studies. Despite the variety of treatments examined, problems treated, or even the presumed motivations of the researchers, the findings from most component studies appear to be relatively homogeneous, regardless of whether the additive and dismantling studies were analyzed separately (see Table 2) or together (see Table 4).

When More Is Less

Although the average effect sizes for the additive studies favored the treatment with the added component over the standard treatment when targeted outcomes were assessed, a review of Table 1 shows that there were a number of studies in which the standard treatment yielded better targeted outcomes at termination or at follow-up. There were 7 additive studies in which the effect size at termination was less than −0.20 at termination. In three of those instances, the effect reversed at follow-up creating the type of scenario that Westen and Morrison (2001) warned of, in which an apparently initially useless or even harmful intervention may have longer-term benefits. For example, Thackwray, Smith, Bodfish, and Meyers (1993) found that behavior therapy appeared to be slightly more effective for treating bulimia than CBT at termination, but at 6-month follow-up, almost twice as many clients in the CBT group were fully recovered (69%) than in the behavior therapy group (38%).

Most of the additive studies that yielded negative effect sizes (i.e., the standard treatment was superior to the treatment with the added component) did not find a significant difference between the two conditions (some of these null findings may have been due to a lack of statistical power). However, among the dismantling studies, there were some studies where the researchers found the partial treatment to be superior to the full treatment (on at least some outcome measures), raising the possibility that some components are not simply unnecessary, but may be iatrogenic. For example, Taylor et al. (2003) found that exposure therapy was superior to EMDR for reducing avoidance and re-experiencing in PTSD. Similarly, although Schmidt et al. (2000) generally found equivalence between CBT with and without breathing retraining for treating panic disorder, there were some indications that the CBT group that did not get breathing retraining had greater improvement on a number of important outcomes, including panic frequency and avoidance. However, it should be noted that in both studies the full treatment (EMDR, CBT with breathing retraining) was superior to the control condition (relaxation and waitlist control, respectively), so perhaps labeling these unnecessary components as iatrogenic may be excessive. Regardless, dismantling designs can be useful for identifying potentially problematic treatment components, but even when there is a sound theoretical basis for questioning the value of component (eye movements, breathing retraining), statistical power issues may still limit the interpretability of these research findings.

Limitations and the Future of Psychotherapy Component Research

The primary limitation of this meta-analysis is that unlike systematic reviews, it is unlikely that our review identified every published psychotherapy component study from the past 30 years. Many psychotherapy studies do not use their designs as index terms, so it is likely that we missed studies. An objection might also be raised that the current meta-analysis mixed apples with oranges. We included treatments for problems as diverse as tension headaches, PTSD, rheumatoid arthritis, and marital discord, and the clients in these studies ranged across the lifespan. Yet, despite the methodological and conceptual heterogeneity of these studies, statistically their results were strikingly homogeneous.

Considering that the strongest findings in the current meta-analysis involved the follow-up data for the additive studies, it is important to acknowledge that the interpretation of follow-up data poses challenges that are not found with termination data. Most of the studies in the meta-analysis had some additional attrition at the follow-up assessment, and it is possible that the participants who benefited the least from therapy were less likely to respond to a request for follow-up data. However, attrition from termination to follow-up was only about 10% of the samples in the additive studies (see Table 3), and the rate of attrition was almost identical in the standard treatment and added component conditions, perhaps mitigating this concern. A second concern is that clients may seek additional treatment during the follow-up interval, and random life events may contribute to the clients' subsequent improvement or relapse. However, these post-treatment events would be likely to attenuate any differences between the standard and added treatment conditions, suggesting that the follow-up effect we found may be a conservative value. Regardless, because of intervening events and the potential for loss of randomization, it is more difficult to make causal attributions for follow-up data than for termination data.

Our findings may also have implications for the development of new treatments. It is not uncommon for treatment developers to add a variety of components to their initial treatment manuals as a way to maximize outcomes in pilot work. Unfortunately, such multi-component interventions may ultimately include unnecessary components that make them difficult to disseminate and compromise their effectiveness in non-academic treatment settings. Considering our findings that additive designs were more likely to yield group differences than dismantling designs, it might be preferable to build interventions by starting with the components that are most likely to be efficacious and then testing additional components using an additive design. Yet, either approach may lead to challenges with respect to statistical power.

Thus, although we agree in principle with Borkovec and Castonguay (1998) that component studies provide elegant designs for identifying the active ingredients in psychotherapy, the findings from the current meta-analysis raise questions about the practical value of any individual component study and underscore the challenges facing researchers who choose to undertake such studies. Although by starting with effective treatments and testing which added components incrementally improve outcomes, additive studies may be an ideal method for advancing the development of psychotherapies, even non-trivial improvements will be difficult to demonstrate. If the current effect sizes for targeted

problems for additive studies of 0.14 at termination and 0.28 at follow-up are accurate estimates, then an additive study examining treatment effects at termination would require 802 clients in each condition to achieve 0.80 power, and one examining follow-up outcomes would require 201 clients in each condition. In contrast, consider that a reformulated antidepressant that yielded an effect size of 0.28 greater than original medication would probably be considered a breakthrough and would be the subject a large scale clinical trials. However such large sample psychotherapy studies are virtually impossible to execute. For example, although behavioral treatments for obsessive compulsive disorder (OCD) are at least as effective as medication, a meta-analysis of psychotherapy and pharmacotherapy for OCD found 15 psychotherapy trials with a total *N* of 705 and 32 pharmacotherapy trials with a total *N* of 3,588 (Eddy, Dutra, Bradley, & Westen, 2004). Not only were there more than twice as many pharmacotherapy trials, but they averaged more than twice the number of participants in each trial. Pharmacotherapy studies are often easier to conduct, and there is greater funding available for them than for psychotherapy studies, raising questions about how researchers will be able to refine psychological treatments. Realistically, it seems unlikely that most additive psychotherapy studies will have the resources required to have sufficient statistical power. Perhaps practice research networks (Borkovec, Echemendia, Ragusea, & Ruiz, 2001) may one day provide an infrastructure for conducting larger scale additive design studies, but for now most additive studies are likely to continue to run the risk of Type II error. Furthermore, future dismantling studies are virtually guaranteed to find small or no differences between the standard treatment and the dismantled condition.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Ahn, H., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology, 48,* 251–257. doi:10.1037/0022-0167.48.3.251

*Appelbaum, K. A., Blanchard, E. B., Nicholson, N. L., Radnitz, C., Kirsch, C., Michultka, D., . . . Dentinger, M. P. (1990). Controlled evaluation of the addition of cognitive strategies to a home-based relaxation protocol for tension headache. *Behavior Therapy, 21,* 293–303. doi:10.1016/S0005-7894(05)80332-6

Barlow, D. H. (2008). *Clinical handbook of psychological disorders* (4th ed.). New York, NY: Guilford Press.

*Barlow, D. H., Craske, M. G., Cerny, J. A., & Klosko, J. S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy, 20,* 261–282. doi:10.1016/S0005-7894(89)80073-5

*Barlow, D. H., Rapee, R. M., & Brown, T. A. (1992). Behavioral treatment of generalized anxiety disorder. *Behavior Therapy, 23,* 551–570. doi:10.1016/S0005-7894(05)80221-7

*Baucom, D. H., Sayers, S. L., & Sher, T. G. (1990). Supplementing behavioral marital therapy with cognitive restructuring and emotional expressiveness training: An outcome investigation. *Journal of Consulting and Clinical Psychology, 58,* 636–645. doi:10.1037/0022-006X.58.5.636

*Bauman, W., & Melnyk, W. T. (1994). A controlled comparison of eye movements and finger tapping in the treatment of test anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 25,* 29–33. doi:10.1016/0005-7916(94)90060-4

*Blanchard, E. B., Appelbaum, K. A., Radnitz, C. L., Michultka, D., Morrill, B., Kirsch, C., . . . Dentinger, M. P. (1990). Placebo-controlled evaluation of abbreviated progressive muscle relaxation and of relaxation combined with cognitive therapy in the treatment of tension headache. *Journal of Consulting and Clinical Psychology, 58,* 210–215. doi:10.1037/0022-006X.58.2.210

Bohart, A. C., & Tallman, K. (1999). *How clients make therapy work: The process of active self-healing*. doi:10.1037/10323-000

Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology, 66,* 136–142. doi:10.1037/0022-006X.66.1.136

*Borkovec, T. D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 61,* 611–619. doi:10.1037/0022-006X.61.4.611

Borkovec, T. D., Echemendia, R. J., Ragusea, S. A., & Ruiz, M. (2001). The Pennsylvania Practice Research Network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology: Science and Practice, 8,* 155–167. doi:10.1093/clipsy.8.2.155

*Borkovec, T. D., Newman, M. G., Pincus, A. L., & Lytle, R. (2002). A component analysis of cognitive-behavioral therapy for generalized anxiety disorder and the role of interpersonal problems. *Journal of Consulting and Clinical Psychology, 70,* 288–298. doi:10.1037/0022-006X.70.2.288

*Bryant, R. A., Moulds, M. L., Guthrie, R. M., Dang, S. T., Mastrodomenico, J., Nixon, R. D. V., . . . Hopwood, S. (2008). A randomized controlled trial of exposure therapy and cognitive restructuring for posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 76,* 695–703. doi:10.1037/a0012616

*Bryant, R. A., Moulds, M. L., Guthrie, R. M., Dang, S. T., & Nixon, R. D. V. (2003). Imaginal exposure alone and imaginal exposure with cognitive restructuring in treatment of posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 71,* 706–712. doi:10.1037/0022-006X.71.4.706

*Bryant, R. A., Moulds, M. L., Guthrie, R. M., & Nixon, R. D. V. (2005). The additive benefit of hypnosis and cognitive-behavioral therapy in treating acute stress disorder. *Journal of Consulting and Clinical Psychology, 73,* 334–340. doi:10.1037/0022-006X.73.2.334

*Butler, G., Cullington, A., Munby, M., Amies, P., & Gelder, M. (1984). Exposure and anxiety management in the treatment of social phobia. *Journal of Consulting and Clinical Psychology, 52,* 642–650. doi:10.1037/0022-006X.52.4.642

Chambless, D. L. (2002). Beware the dodo bird: The dangers of overgeneralization. *Clinical Psychology: Science and Practice, 9,* 13–16. doi:10.1093/clipsy.9.1.13

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Craske, M. G., Brown, T. A., & Barlow, D. H. (1991). Behavioral treatment of panic disorder: A two-year follow-up. *Behavior Therapy, 22,* 289–304. doi:10.1016/S0005-7894(05)80367-3

Crits-Christoph, P. (1997). Limitations of the dodo bird verdict and the role of clinical trials in psychotherapy research: Comment on Wampold et al. (1997). *Psychological Bulletin, 122,* 216–220. doi:10.1037/0033-2909.122.3.216

*Cusack, K., & Spates, R. C. (1999). The cognitive dismantling of eye movement desensitization and reprocessing (EMDR) treatment of posttraumatic stress disorder (PTSD). *Journal of Anxiety Disorders, 13,* 87–99. doi:10.1016/S0887-6185(98)00041-3

*Dadds, M. R., & McHugh, T. A. (1992). Social support and treatment outcome in behavioral family therapy for child conduct problems. *Journal of Consulting and Clinical Psychology, 60,* 252–259. doi:10.1037/0022-006X.60.2.252

*Deffenbacher, J. L., Filetti, L. B., Lynch, R. S., Dahlen, E. R., & Oetting, E. R. (2002). Cognitive-behavioral treatment of high anger drivers. *Behaviour Research and Therapy, 40,* 895–910. doi:10.1016/S0005-7967(01)00067-5

*Deffenbacher, J. L., & Stark, R. S. (1992). Relaxation and cognitive-relaxation treatments of general anger. *Journal of Counseling Psychology, 39,* 158–167. doi:10.1037/0022-0167.39.2.158

*de Jong, P. J., Vorage, I., & van den Hout, M. A. (2000). Counterconditioning in the treatment of spider phobia: Effects of disgust, fear and valence. *Behaviour Research and Therapy, 38,* 1055–1069. doi:10.1016/S0005-7967(99)00135-7

*Devilly, G. J., Spence, S. H., & Rapee, R. M. (1998). Statistical and reliable change with eye movement desensitization and reprocessing: Treating trauma within a veteran population. *Behavior Therapy, 29,* 435–455. doi:10.1016/S0005-7894(98)80042-7

*de Zwaan, M., Mitchell, J. E., Crosby, R. D., Mussell, M. P., Raymond, N. C., Specker, S. M., & Seim, H. C. (2005). Short-term cognitive behavioral treatment does not improve outcome of a comprehensive very-low-calorie diet program in obese women with binge eating disorder. *Behavior Therapy, 36,* 89–99. doi:10.1016/S0005-7894(05)80057-7

Dimidjian, S., Martell, C. R., Addis, M. E., & Herman-Dunn, R. (2008). Behavioral activation for depression. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders* (4th ed., pp. 328–364). New York, NY: Guilford Press.

*Dunn, T. M., Schwarts, M., Hatfield, R. W., & Wiegele, M. (1996). Measuring effectiveness of eye movement desensitization and reprocessing (EMDR) in non-clinical anxiety: A multi-subject, yoked control design. *Journal of Behavior Therapy and Experimental Psychiatry, 27,* 231–239. doi:10.1016/S0005-7916(96)00034-1

Duval, S. J., & Tweedie, R. L. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89–98.

Eddy, K. T., Dutra, L., Bradley, R., & Westen, D. (2004). A multidimensional meta-analysis of psychotherapy and pharmacotherapy for obsessive-compulsive disorder. *Clinical Psychology Review, 24,* 1011–1030. doi:10.1016/j.cpr.2004.08.004

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629–634. doi:10.1136/bmj.315.7109.629

*Emmelkamp, P. M. G., & Beens, H. (1991). Cognitive therapy with obsessive compulsive disorder: A comparative evaluation. *Behaviour Research and Therapy, 29,* 293–300. doi:10.1016/0005-7967(91)90120-R

*Feske, U., & Goldstein, A. J. (1997). Eye movement desensitization and reprocessing treatment for panic disorder: A controlled outcome and partial dismantling study. *Journal of Consulting and Clinical Psychology, 65,* 1026–1035. doi:10.1037/0022-006X.65.6.1026

*Flessner, C. A., Miltenberger, R. G., Egemo, K., Kelso, P., Jostad, C., Johnson, B., . . . Neighbors, C. (2005). An evaluation of the social support component of simplified habit reversal. *Behavior Therapy, 36,* 35–42. doi:10.1016/S0005-7894(05)80052-8

*Foa, E. B., Hembree E. A., Cahill, S. P., Rauch, A. M., Riggs, D. S., Feeny, N. C., & Yadin, E. (2005). Randomized trial of prolonged exposure for posttraumatic stress disorder with and without cognitive restructuring: Outcome at academic and community clinics. *Journal of Consulting and Clinical Psychology, 73,* 953–964. doi:10.1037/0022-006X.73.5.953

*Foa, E. B., & Rauch, S. A. M. (2004). Cognitive changes during prolonged exposure versus prolonged exposure plus cognitive restructuring in female assault survivors with posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 72,* 879–884. doi:10.1037/0022-006X.72.5.879

*Foa, E. B., Steketee, G., Grayson, J. B., Turner, R. M., & Latimer, P. R. (1984). Deliberate exposure and blocking of obsessive-compulsive rituals: Immediate and long-term effects. *Behavior Therapy, 15,* 450–472. doi:10.1016/S0005-7894(84)80049-0

*Foley, T., & Spates, C. R. (1995). Eye movement desensitization of public-speaking anxiety: A partial dismantling. *Journal of Behavior Therapy and Experimental Psychiatry, 26,* 321–329. doi:10.1016/0005-7916(95)00048-8

*Gosselin, P., & Matthews, W. J. (1995). Eye movement desensitization and reprocessing in the treatment of test anxiety: A study of the effects of expectancy and eye movement. *Journal of Behavior Therapy and Experimental Psychiatry, 26,* 331–337. doi:10.1016/0005-7916(95)00038-0

*Grunes, M. S., Neziroglu, F., & McKay, D. (2001). Family involvement in the behavioral treatment of obsessive-compulsive disorder: A preliminary investigation. *Behavior Therapy, 32,* 803–820. doi:10.1016/S0005-7894(01)80022-8

*Halford, W. K., Sanders, M. R., & Behrens, B. C. (1993). A comparison of the generalization of behavioral marital therapy and enhanced behavioral marital therapy. *Journal of Consulting and Clinical Psychology, 61,* 51–60. doi:10.1037/0022-006X.61.1.51

Herbert, J. D., Lilienfeld, S. O., Lohr, J. M., Montgomery, R. W., O'Donohue, W. T., Rosen, G. M., & Tolin, D. F. (2000). Science and pseudoscience in the development of eye movement desensitization and reprocessing: Implications for clinical psychology. *Clinical Psychology Review, 20,* 945–971. doi:10.1016/S0272-7358(99)00017-3

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal, 327,* 557–560. doi:10.1136/bmj.327.7414.557

Higgins, J. P. T., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society, 172,* 137–159. doi:10.1111/j.1467-985X.2008.00552.x

*Hiss, H., Foa, E. B., & Kozak, M. J. (1994). Relapse prevention for treatment of obsessive-compulsive disorder. *Journal of Consulting and Clinical Psychology, 62,* 801–808. doi:10.1037/0022-006X.62.4.801

*Hope, D. A., Heimberg, R. G., & Bruch, M. A. (1995). Dismantling cognitive-behavioural group therapy for social phobia. *Behaviour Research and Therapy, 33,* 637–650. doi:10.1016/0005-7967(95)00013-N

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist, 41,* 159–164. doi:10.1037/0003-066X.41.2.159

Hunsley, J., & Di Giulio, G. (2002). Dodo bird, phoenix, or urban legend? The question of psychotherapy equivalence. *The Scientific Review of Mental Health Practice, 1,* 11–22.

*Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., & Gollan, J. K. (1996). A component analysis of cognitive-behavioral treatment of depression. *Journal of Consulting and Clinical Psychology, 64,* 295–304. doi:10.1037/0022-006X.64.2.295

*Kazdin, A. E., & Whitley, M. K. (2003). Treatment of parental stress to enhance therapeutic change among children referred for aggressive and antisocial behavior. *Journal of Consulting and Clinical Psychology, 71,* 504–515. doi:10.1037/0022-006X.71.3.504

*Koch, E. I., Spates, C. R., & Himle, J. A. (2004). Comparison of behavioral and cognitive-behavioral one-session exposure treatments for small animal phobias. *Behaviour Research and Therapy, 42,* 1483–1504. doi:10.1016/j.brat.2003.10.005

Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 143–189). New York, NY: Wiley.

Lang, P. J. (1969). The mechanics of desensitization and the laboratory study of human fear. In F. M. Franks (Ed.), *Behavior therapy: Appraisal and status* (pp. 161–162). New York, NY: McGraw-Hill Book Company.

Leykin, Y., & DeRubeis, R. J. (2009). Allegiance in psychotherapy outcome research: Separating association from bias. *Clinical Psychology:*

Science and Practice, 16, 54–65. doi:10.1111/j.1468-2850.2009.01143.x

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavior therapy? *Clinical Psychology Review, 27,* 173–187. doi:10.1016/j.cpr.2006.08.001

Luborsky, L., Diguer, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., . . . Schweizer, E. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice, 6,* 95–106. doi:10.1093/clipsy/6.1.95

Luborsky, L., Rosenthal, R., Diguer, L., Andrusyana, T. P., Berman, J. S., Levitt, J. T., . . . Krause, E. D. (2002). The dodo bird verdict is alive and well-mostly. *Clinical Psychology: Science and Practice, 9,* 2–12. doi:10.1093/clipsy/9.1.2

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "everyone has won and all must have prizes"? *Archives of General Psychiatry, 32,* 995–1008. doi:10.1001/archpsyc.1975.01760260059004

*Marks, I., Lovell, K., Noshirvani, H., Livanou, M., & Thrasher, S. (1998). Treatment of posttraumatic stress disorder by exposure and/or cognitive restructuring. *Archives of General Psychiatry, 55,* 317–325. doi:10.1001/archpsyc.55.4.317

*Mattick, R. P., & Peters, L. (1988). Treatment of severe social phobia: Effects of guided exposure with and without cognitive restructuring. *Journal of Consulting and Clinical Psychology, 56,* 251–260. doi:10.1037/0022-006X.56.2.251

*Mattick, R. P., Peters, L., & Clarke, J. C. (1989). Exposure and cognitive restructuring for social phobia: A controlled study. *Behavior Therapy, 20,* 3–23. doi:10.1016/S0005-7894(89)80115-7

*McKay, J. R., Lynch, K. G., Coviello, D., Morrison, R., Cary, M. S., Skalina, L., & Plebani, J. (2010). Randomized trial of continuing care enhancements for cocaine-dependent patients following initial engagement. *Journal of Consulting and Clinical Psychology, 78,* 111–120. doi:10.1037/a0018139

*Mehta, M. (1990). A comparative study of family-based and patient-based behavioural management in obsessive-compulsive disorder. *The British Journal of Psychiatry, 157,* 133–135. doi:10.1192/bjp.157.1.133

Munder, T., Gerger, H., Trelle, S., & Barth, J. (2011). Testing the allegiance bias hypothesis: A meta-analysis. *Psychotherapy Research, 21,* 670–684. doi:10.1080/10503307.2011.602752

*Nezu, A. M., & Perri, M. G. (1989). Social problem-solving therapy for unipolar depression: An initial dismantling investigation. *Journal of Consulting and Clinical Psychology, 57,* 408–413. doi:10.1037/0022-006X.57.3.408

*Nicholas, M. K., Wilson, P. H., & Goyen, J. (1991). Operant-behavioural and cognitive-behavioral treatment for chronic low back pain. *Behaviour Research and Therapy, 29,* 225–238. doi:10.1016/0005-7967(91)90112-G

*Öst, L.-G., Fellenius, J., & Sterner, U. (1991). Applied tension, exposure in vivo, and tension-only in the treatment of blood phobia. *Behaviour Research and Therapy, 29,* 561–574. doi:10.1016/0005-7967(91)90006-O

*Öst, L.-G., Thulin, U., & Ramnero, J. (2004). Cognitive-behavior therapy vs. exposure in vivo in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy, 42,* 1105–1127. doi:10.1016/j.brat.2003.07.004

*Paunovic, N., & Öst, L. (2001). Cognitive-behavior therapy vs. exposure therapy in the treatment of PTSD in refugees. *Behaviour Research and Therapy, 39,* 1183–1197. doi:10.1016/S0005-7967(00)00093-0

*Petry, N. M., Weinstock, J., Ledgerwood, D. M., & Morasco, B. (2008). A randomized trial of brief interventions for problem and pathological gamblers. *Journal of Consulting and Clinical Psychology, 76,* 318–328. doi:10.1037/0022-006X.76.2.318

*Pitman, R. K., Orr, S. P., Altman, B., Longpre, R. E., Poire, R. E., & Macklin, M. L. (1996). Emotional processing during eye movement desensitization and reprocessing therapy of Vietnam veterans with chronic posttraumatic stress disorder. *Comprehensive Psychiatry, 37,* 419–429. doi:10.1016/S0010-440X(96)90025-5

*Porzelius, L. K., Houston, C., Smith, M., Arfken, C., & Fisher, E., Jr. (1995). Comparison of a standard behavioral weight loss treatment and a binge eating weight loss treatment. *Behavior Therapy, 26,* 119–134. doi:10.1016/S0005-7894(05)80086-3

*Propst, L. R., Ostrom, R., Watkins, P., Dean, T., & Mashburn, D. (1992). Comparative efficacy of religious and nonreligious cognitive-behavioral therapy for the treatment of clinical depression in religious individuals. *Journal of Consulting and Clinical Psychology, 60,* 94–103. doi:10.1037/0022-006X.60.1.94

*Radojevic, V., Nicassio, P. M., & Weisman, M. H. (1992). Behavioral intervention with and without family support for rheumatoid arthritis. *Behavior Therapy, 23,* 13–30. doi:10.1016/S0005-7894(05)80305-3

*Rehm, L. P. (1981). An evaluation of major components in a self-control therapy program for depression. *Behavior Modification, 5,* 459–489. doi:10.1177/014544558154002

Rehm, L. P. (2009). Looking at the components of treatment for depression. *Clinical Psychology: Science and Practice, 16,* 412–415. doi:10.1111/j.1468-2850.2009.01179.x

*Resick, P. A., Galoviski, T. E., Uhlmansiek, M. O., Scher, C. D., Clum, G. A., & Young-Xu, Y. (2008). A randomized clinical trial to dismantle components of cognitive processing therapy for posttraumatic stress disorder in female victims of interpersonal violence. *Journal of Consulting and Clinical Psychology, 76,* 243–258. doi:10.1037/0022-006X.76.2.243

*Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ: British Medical Journal, 342,* 964–967.

*Roehrig, M., Thompson, J. K., Brannick, M., & van den Berg, P. (2006). Dissonance-based eating disorder prevention program: A preliminary dismantling investigation. *International Journal of Eating Disorders, 39,* 1–10. doi:10.1002/eat.20217

*Rohan, K. J., Roecklein, K. A., Lindsey, K. T., Johnson, L. G., Lippy, R. D., Lacey, T. J., & Barton, F. B. (2007). A randomized controlled trial of cognitive-behavioral therapy, light therapy, and their combination for seasonal affective disorder. *Journal of Consulting and Clinical Psychology, 75,* 489–500. doi:10.1037/0022-006X.75.3.489

*Rosen, J. C., Cado, S., Silberg, N. T., Srebnik, D., & Wendt, S. (1990). Cognitive behavior therapy with and without size perception training for women with body image disturbance. *Behavior Therapy, 21,* 481–498. doi:10.1016/S0005-7894(05)80360-0

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74,* 166–169. doi:10.1037/0022-0663.74.2.166

*Sanders, M. R., & McFarland, M. (2000). Treatment of depressed mothers with disruptive children: A controlled evaluation of cognitive behavioral family intervention. *Behavior Therapy, 31,* 89–112. doi:10.1016/S0005-7894(00)80006-4

*Sanderson, A., & Carpenter, R. (1992). Eye movement desensitization versus image confrontation: A single-session crossover study of 58 phobic subjects. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 269–275. doi:10.1016/0005-7916(92)90049-O

*Schmidt, N. B., Woolaway-Bickel, K., Trakowski, J., Santiago, H., Storey, J., Koselka, M., & Cook, J. (2000). Dismantling cognitive-behavioral treatment for panic disorder: Questioning the utility of breathing retraining. *Journal of Consulting and Clinical Psychology, 68,* 417–424. doi:10.1037/0022-006X.68.3.417

*Schmiege, S. J., Broaddus, M. R., Levin, M., & Bryan, A. D. (2009). Randomized trial of group interventions to reduce HIV/STD risk and

change theoretical mediators among detained adolescents. *Journal of Consulting and Clinical Psychology, 77,* 38–50. doi:10.1037/a0014513

Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychological treatments. *The Clinical Psychologist, 48,* 3–23.

*Taylor, S., Thordarson, D. S., Fedoroff, I. C., Maxfield, L., & Lovell, K. (2003). Comparative efficacy, speed, and adverse effects of three PTSD treatments: Exposure therapy, EMDR, and relaxation training. *Journal of Consulting and Clinical Psychology, 71,* 330–338. doi:10.1037/0022-006X.71.2.330

*Taylor, S., Woody, S., Koch, W. J., McLean, P., Paterson, R. J., & Anderson, K. W. (1997). Cognitive restructuring in the treatment of social phobia: Mode of action. *Behavior Modification, 21,* 487–511. doi:10.1177/01454455970214006

*Thackwray, D. E., Smith, M. C., Bodfish, J. W., & Meyers, A. W. (1993). A comparison of behavioral and cognitive-behavioral interventions for bulimia nervosa. *Journal of Consulting and Clinical Psychology, 61,* 639–645. doi:10.1037/0022-006X.61.4.639

*Tolin, D. F. (2010). Is cognitive–behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review, 30,* 710–720. doi:10.1016/j.cpr.2010.05.003

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

*Walters, S. T., Vader, A. M., Harris, T., Field, C. A., & Jouriles, E. N. (2009). Dismantling motivational interviewing and feedback for college drinkers: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 77,* 64–73. doi:10.1037/a0014472

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes". *Psychological Bulletin, 122,* 203–215. doi:10.1037/0033-2909.122.3.203

*Webster-Stratton, C. (1994). Advancing videotape parent training: A comparison study. *Journal of Consulting and Clinical Psychology, 62,* 583–593. doi:10.1037/0022-006X.62.3.583

Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology, 69,* 875–899. doi:10.1037/0022-006X.69.6.875

*Williams, S. L., & Falbo, J. (1996). Cognitive and performance-based treatments for panic attacks in people with varying degrees of agoraphobic disability. *Behaviour Research and Therapy, 34,* 253–264. doi:10.1016/0005-7967(95)00063-1

Received December 6, 2011
Revision received March 15, 2013
Accepted April 10, 2013 ■